



Statistical analysis of PET radiomic features and metabolomic data: prediction of triple-negative breast cancer

Fanny Orlhac, Olivier Humbert, Thierry Pourcher, Lun Jing, Jean-Marie Guigonis, Jacques Darcourt, Nicholas Ayache, Charles Bouveyron

► To cite this version:

Fanny Orlhac, Olivier Humbert, Thierry Pourcher, Lun Jing, Jean-Marie Guigonis, et al.. Statistical analysis of PET radiomic features and metabolomic data: prediction of triple-negative breast cancer. SNMMI Annual Meeting, Jun 2018, Philadelphia, United States. pp.1755. hal-01759330

HAL Id: hal-01759330

<https://hal.science/hal-01759330>

Submitted on 5 Apr 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Statistical analysis of PET radiomic features and metabolomic data: prediction of triple-negative breast cancer

Fanny Orlhac¹, Olivier Humbert^{2,3,4}, Thierry Pourcher^{3,4}, Lun Jing^{3,4}, Jean-Marie Guignonis^{3,4}, Jacques Darcourt^{2,3,4}, Nicholas Ayache¹, Charles Bouveyron^{1,5}.

1: Epione, Inria Sophia-Antipolis, Valbonne, France, 2: Department of Nuclear Medicine, Centre Antoine Lacassagne/UCA, Nice, France, 3 : Laboratoire TIRO, UMRE 4320 CEA, Nice, France, 4 : Université Côte d'Azur, Nice, France, 5 : Laboratoire JAD, UMR CNRS 7135, Université Côte d'Azur, Nice, France.

Objective: The characterization of tumor heterogeneity using radiomic features from PET images is gaining interest in the context of precision medicine. Yet, the relationship between radiomic features and biological characteristics of lesions needs to be clarified. To this end, we studied the relationship between PET radiomic features and metabolomic data in breast cancer and we investigated their ability to predict the immunohistochemical tumor subtypes.

Methods: 26 patients with a breast cancer underwent a pre-treatment FDG PET/CT scan on a Biograph PET/CT scanner (Siemens). In each patient, the primary lesion was segmented using a threshold equal to 40% of SUVmax. In each volume of interest, we computed 43 radiomic features using the LIFEx software (absolute discretization: 64 gray-levels between 0 and 20 SUV units) including SUVmax, SUVmean, Metabolic Volume (MV), TLG, 7 histogram indices, 2 shape features and 30 textural indices. Based on the resected tissues, we used a mass spectrometer to analyze the expression of 1500 metabolites listed in the Human metabolome database. Spearman correlation coefficients (R) between each radiomic feature and each metabolite were studied. We investigated the ability of radiomic features and metabolomic data to identify triple-negative breast cancer (TNBC). We compared the performance of logistic regression for SUVmax, MV and SUVmax+MV with those of 5 statistical methods for radiomic features and metabolomic data separately: linear discriminant analysis (LDA), partial least squares discriminant analysis (PLSDA), orthogonal partial least squares (OPLS), high dimensional discriminant analysis (HDDA, [1]) and globally sparse HDDA (gsHDDA, [2]). This procedure was repeated 25 times with 16 randomly selected patients for the training set and 10 patients for the validation set. The accuracy of the TNBC identification was measured on the validation set using the Youden index (sensitivity + specificity - 1).

Results: In our cohort, 7 women had TNBC. The mean correlation coefficient, in absolute value, between radiomic and metabolomic features was equal to 0.20 ± 0.14 (range: [0-0.81]). Only 3% of pairwise correlations were higher than 0.50 (in absolute value). Twenty out of 43 radiomic features were moderately correlated with at least 50 metabolites ($|R| \geq 0.50$) including SUVs and MV. With the logistic regression, the best performance for TNBC identification was obtained for SUVmax with a Youden index equal to 0.29 ± 0.34 . Using different statistical methods, the Youden indices ranged between 0.18 and 0.34 based on metabolomic data and between -0.12 and 0.50 from radiomic features. The best performance for TNBC identification was obtained for HDDA (Youden = 0.50 ± 0.35) and gsHDDA (Youden = 0.49 ± 0.34) based on radiomic features and these results outperformed those obtained with SUVmax (p-values of Wilcoxon test = 0.03). One of the advantages of gsHDDA compared to other methods is that the model was built based on selected features. The study of these features showed a moderate correlation ($|R| = 0.21 \pm 0.15$, range: [0-0.71]) between 10 key radiomic features and 601 key metabolites (features selected by gsHDDA in at least 50% of the tests), suggesting that the combination between these two sources of information could improve the identification of TNBC.

Conclusion: In breast lesions, we demonstrated a poor to moderate correlation between PET radiomic features and metabolomic data. However, the two types of data allow the identification of TNBC with similar performances. Additional breast cancer patients are currently being included to validate these results on a large cohort and the joint analysis of radiomic features and metabolomic data will be investigated in order to take advantage of the complementarity of data and enhance classification performance.